

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»**

ФІЗИКО- ТЕХНІЧНИЙ ІНСТИТУТ

Кафедра інформаційної безпеки

КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

з дисципліни

Криптографія

З теми: « Експериментальна оцінка ентропії на символ джерела відкритого тексту »

Перевірила:

Селюх П.В

Виконав студент групи ФБ-92

Андрієвич Дмитро Юрійович

Мета роботи: Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела.

Завдання:

0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку H_1 та H_2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення H_1 та H_2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення H_1 та H_2 на тому ж тексті, в якому вилучено всі пробіли.
2. За допомогою програми CoolPinkProgram оцінити значення $H(10)$, $H(20)$, $H(30)$.
3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела

1.

Код програми знаходиться у файлах **main.py** та **text_processing.py**.

Результати:

H_1 with whitespaces: 4.376013208231846

H_2 with whitespaces: 2.9039857423897684

H_1 without whitespaces: 4.4507435574840315

H_2 without whitespaces: 3.3499226595401437

Character frequency without whitespaces

	Frequency
о	0,110649
а	0,08638
е	0,081299
и	0,068439
н	0,064124
т	0,0603
л	0,052543
с	0,050689
р	0,047465
в	0,047094
к	0,036579
у	0,030055
м	0,030013
п	0,028431
д	0,02811
г	0,019153
я	0,019069
ь	0,018088
з	0,017777
ы	0,017297
ч	0,015901
б	0,015486
й	0,011723
ж	0,009011
ш	0,008731
х	0,008228
ю	0,00528
щ	0,003538
ц	0,003317
э	0,003072
ф	0,002159

Character frequency with whitespaces

	Frequency
	0,158498
о	0,093112
а	0,072689
е	0,068413
и	0,057592
н	0,05396
т	0,050742
л	0,044215
с	0,042655
р	0,039942
в	0,03963
к	0,030781
у	0,025291
м	0,025256
п	0,023925
д	0,023655
г	0,016117
я	0,016047
ь	0,015221
з	0,014959
ы	0,014556
ч	0,013381
б	0,013032
й	0,009865
ж	0,007582
ш	0,007347
х	0,006924
ю	0,004443
щ	0,002977
ц	0,002792
э	0,002585
ф	0,001816

Bigram frequency with whitespaces (**bigram_frequency_with_whitespaces.xlsx**)

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ы	ь	э	ю	я			
а	1,27E-05	0,000556	0,002468	0,000902	0,002739	0,001103	0,001088	0,003873	0,000293	0,000577	0,000456	0,000864	0,003104	0,005127	3,24E-05	0,000865	0,0004363	0,000403	0,0004899	7,6E-05	0,000187	0,0001126	7,18E-05	0,000922	0,000791	0,000342	0	0	1,41E-05	0,000733	0,00212			
б	0,000954	1,13E-05	2,11E-05	1,83E-05	7,04E-06	0,002319	1,13E-05	1,41E-06	0,00066	0	0,000159	0,000568	4,5E-05	0,000317	0,00194	0	0,001071	8,16E-05	4,22E-06	0,0001216	0	2,81E-05	1,41E-06	5,63E-06	7,04E-06	0,000129	0,002742	0,000179	0	1,69E-05	0,000324			
в	0,006295	1,13E-05	0,000111	1,97E-05	0,00027	0,005621	0	0,000414	0,003523	0	0,000232	0,000585	0,000117	0,000881	0,006305	0,000217	0,000612	0,000212	0,000288	0,000651	0	3,94E-05	3,66E-05	0,000108	0,000954	1,69E-05	0,00261	0,000484	0	0	0,000256			
г	0,002139	0	2,81E-06	0	0,000902	0,00048	0	1,41E-06	0,00063	0	7,46E-05	0,001522	1,13E-05	0,00039	0,007613	1,41E-06	0,00128	8,44E-06	2,11E-05	0,000483	0	3,94E-05	0	4,92E-05	1,41E-06	0	0	0	0	0	0			
д	0,004191	7,04E-05	0,000916	1,13E-05	1,13E-05	0,003858	5,21E-05	9,85E-06	0,001984	0	0,000301	0,000421	0,000129	0,001929	0,003259	0,000101	0,000965	0,000357	0,000198	0,001458	0	3,38E-05	0,000211	6,47E-05	0,00016	0	0,000487	0,000581	0	9,85E-06	0,000412			
е	0,000131	0,000862	0,002088	0,002936	0,002813	0,001292	0,000776	0,001384	0,0002	0,001856	0,001437	0,005929	0,00365	0,000624	0,000194	0,001186	0,006484	0,000468	0,005192	0,00016	0,000118	0,000397	0,000331	0,001062	0,000651	0,00058	0	0	0,000179	0,000172	0	0		
ж	0,001318	2,39E-05	0	1,41E-05	0,000718	0,002962	9,85E-06	0	0,001131	0	0,000141	4,08E-05	8,44E-06	0,00063	5,35E-05	0	4,22E-06	2,11E-05	2,81E-06	0,000186	0	0	0	3,52E-05	0	0	0	1,83E-05	0	0	0	0		
з	0,005469	0,00017	0,00109	0,000273	0,000864	0,000629	8,72E-05	8,72E-05	0,000491	0	0,000113	0,000345	0,000235	0,001382	0,000649	3,94E-05	0,000259	1,27E-05	4,22E-06	0,000425	0	0	0	1,41E-05	2,81E-06	0	0,000383	0,000152	0	1,41E-06	0,0002			
и	0,000229	0,000563	0,002762	0,000492	0,001486	0,001655	0,000248	0,002194	0,000483	0,001621	0,002676	0,006139	0,002535	0,00324	0,000404	0,000297	0,001107	0,002905	0,00409	0,000121	7,74E-05	0,001441	0,000972	0,001712	0,000497	0,000197	0	0	2,81E-06	0,000245	0,000992			
й	0	1,41E-05	0	1,13E-05	0,000101	1,41E-06	0	0	0	0	0,000131	1,55E-05	3,94E-05	0,000386	1,13E-05	1,41E-06	0	0,000371	0,000387	0	2,81E-06	0	8,3E-05	0,000166	0,000172	1,83E-05	0	0	0	1,41E-05	0	0		
к	0,007083	1,41E-06	0,00054	2,81E-06	0	0,000867	4,78E-05	1,55E-05	0,002887	0	2,25E-05	0,000808	2,81E-06	0,000684	0,008418	0	0,002106	0,000142	0,000701	0,002184	0	2,81E-06	2,53E-05	2,81E-06	1,13E-05	0	0	0	0	0	0			
л	0,007477	7,88E-05	2,39E-05	0,000114	6,61E-05	0,000452	0,000212	2,95E-05	0,006239	0	0,000397	0,000428	5,35E-05	0,000046	0,006213	5,77E-05	0	0,001966	0,000196	0,001301	1,13E-05	5,63E-06	0	0,000132	1,41E-05	1,41E-06	0,000858	0,003503	1,41E-06	0,000677	0,00132			
м	0,003579	1,13E-05	6,89E-05	7,6E-05	7,04E-06	0,003111	0	0	0,002804	0	0,000108	0,000184	8,44E-05	0,001278	0,002905	0,000173	9,43E-05	0,00027	7,6E-05	0,001725	4,22E-06	0	7,04E-06	2,67E-05	1,13E-05	2,81E-06	0,000698	4,5E-05	2,81E-06	0	0,000516			
н	0,009154	5,63E-06	2,95E-05	9,43E-05	0,000854	0,008806	1,13E-05	5,63E-05	0,007029	0	0,000259	0	0	0,003102	0,009727	0	3,94E-05	0,000497	0,000816	0,003736	6,89E-05	0	0,000304	0,000167	4,22E-06	0,000173	0,003195	0,00092	0	0,000158	0,001399			
о	1,13E-05	0,002983	0,006955	0,000435	0,004501	0,001529	0,001565	0,001482	0,000663	0,003814	0,002554	0,006271	0,005669	0,005369	0,000194	0,001356	0,006171	0,005573	0,000484	5,63E-05	0,000307	0,000484	0,000121	0,00171	0,000772	0,000221	0	0	0,000141	0,0004	0,000504			
п	0,001667	0	0	0	0,000215	0	0	0,001255	0	9E-05	0,000896	0	0,000225	0,006174	6,19E-05	0,006498	3,38E-05	0,000152	0,000799	0	0	4,22E-06	2,11E-05	1,69E-05	0	0,000331	8,72E-05	0	2,95E-05	0,00038	0	0		
р	0,000754	6,75E-05	0,000487	0,000103	0,000257	0,00561	0,000228	4,64E-05	0,005569	0	0,000397	0,000256	0,000197	0,000957	0,00793	6,33E-05	5,91E-05	0,00016	0,000924	0,002493	4,22E-06	0,000211	6,61E-05	0,000184	0,000447	4,08E-05	0,001335	0,000623	0	0,00025	0,001016			
с	0,001667	7,18E-05	0,001479	2,67E-05	0,00025	0,002949	3,8E-05	1,83E-05	0,001801	0	0,000389	0,003469	0,000594	0,000862	0,002604	0,001612	0,000194	0,000919	0,002947	0,00075	2,11E-05	0,000131	7,88E-05	0,000318	0,000134	2,81E-06	0,000321	0,003454	1,41E-06	9,43E-05	0,003751			
т	0,005577	2,39E-05	0,001977	2,81E-06	0,000152	0,00571	0	4,22E-06	0,003432	0	0,000722	0,000212	2,81E-05	0,001248	0,013248	8,16E-05	0,000853	0,000787	5,35E-05	0,001968	5,63E-05	1,27E-05	1,69E-05	0,000357	5,63E-06	0,001408	0,00454	0	7,04E-05	0,000446	0	0		
у	0,000103	0,000636	0,000718	0,00093	0,001801	0,000174	0,001195	0,000279	2,11E-05	0,000146	0,000948	0,002186	0,00093	0,000499	4,22E-06	0,000618	0,000968	0,001204	0,000192	4,22E-06	0,000196	0,000833	8,44E-06	0,000675	0,000692	0,00025	0	0	1,55E-05	0,000995	8,86E-05	0	0	
ф	0,000273	1,41E-06	0	0	0,000504	0	0	0,000335	0	0	8,86E-05	0	1,41E-06	0,000245	0	0,000186	2,25E-05	1,69E-05	5,21E-05	1,13E-05	0	0	0	0	0	0	0	3,1E-05	1,41E-05	0	0	0	0	
х	0,000567	1,41E-06	0,000166	5,77E-05	0	6,33E-05	0	0	0,000314	0	7,04E-06	0,000167	5,77E-05	0,000332	0,00178	0	0,000197	4,5E-05	2,11E-05	0,000184	0	2,81E-06	0	2,39E-05	0	0	7,04E-06	2,81E-06	0	0	0	0	0	
ц	0,000632	1,41E-06	6,33E-05	0	0	0,000832	0	0	0,000284	0	1,41E-05	5,63E-06	0	0	0,000257	0	0	2,81E-06	0,000155	0	0	2,81E-06	0	0	0	0	0	0,000156	0	0	0	0	0	
ч	0,002244	0	1,41E-06	0	0	0,002079	0	0	0,001551	0	0,000491	2,53E-05	4,22E-06	0,000778	6,33E-05	0	6,61E-05	0	0,003048	0,000478	0	0	0	5,21E-05	0	0	0	0,000255	0	0	0	0	0	
ш	0,000102	0	6,47E-05	0	0	0,002058	0	0	0,001805	0	0,000391	0,000429	7,04E-06	0,000381	0,002442	4,78E-05	5,63E-06	0	9,57E-05	0,000388	0	0	0	0	0	0	0	0	0,000336	0	0	0	0	0
щ	0,000442	0	0	0	0,001247	0	0	0,001029	0	0	0	0	0	0	4,59E-05	4,22E-06	0	2,81E-06	0	0	0,000144	0	0	0	0	0	0	0	2,11E-05	0	0	0	0	0
ы	0	0,000256	0,000833	0,000101	9,85E-05	0,000895	7,32E-05	7,46E-05	2,81E-06	0,001783	0,000201	0,001914	0,00124	0,000152	0	0,000136	0,000252	0,000664	0,000547	0	0	0,000837	1,55E-05	0,000152	0,000499	7,04E-06	0	0	0	0	1,55E-05	0	0	
ь	0	3,94E-05	3,94E-05	7,46E-05	0,000108	0,000885	1,41E-06	0,000121	5,91E-05	0	0,000905	0	0,000225	0,001247	7,04E-06	1,13E-05	1,41E-06	0,000704	0,000245	0	1,27E-05	0	0,000118	5,91E-05	0,000322	9,85E-06	0	0	4,22E-06	0,000329	0,000415	0	0	0
э	0	0	0	2,42E-06	0	0	0	4,22E-06	0	0	9,85E-06	3,1E-05	2,11E-05	9,85E-06	1,69E-05	0	4,22E-06	1,13E-05	1,41E-05	0,002413	0	7,04E-06	2,39E-05	0	0	0	0	0	0	0	0	0	0	
ю	0	0,000167	1,69E-05	8,44E-06	0,000234	0	5,63E-06	1,41E-06	0	0	3,66E-05	2,67E-05	6,05E-05	3,94E-05	0	1,41E-06	4,64E-05	0,00017	0,000288	0	1,41E-06	6,61E-05	1,69E-05	0,000114	3,24E-05	0,00047	0	0	5,77E-05	1,41E-06	0	0	0	
я	0	2,25E-05	0,000446	9,71E-05	0,000552	9,29E-05	0,00012	0,000259	2,67E-05	4,78E-05	0,0002	0,000812	0,000308	0,000547	0	5,77E-05	9,29E-05	0,000758	0,00124	1,13E-05	7,04E-06	0,000148	4,22E-05	0,000156	2,39E-05	0,00037	0	0	0,000142	0,000131	0	0	0	

Bigram frequency without whitespaces (bigram_frequency.xlsx)

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ы	ь	э	ю	я		
а	1,5E-05	0,00066	0,002933	0,001072	0,00266	0,001311	0,001292	0,004603	0,000348	0,000686	0,005058	0,010268	0,003688	0,006093	3,85E-05	0,001028	0,005185	0,004789	0,005822	9,03E-05	0,000222	0,001338	8,53E-05	0,001095	0,00094	0,000406	0	0	1,67E-05	0,000871	0,00252		
б	0,001134	1,34E-05	2,51E-05	2,17E-05	8,36E-06	0,002755	1,34E-05	1,67E-06	0,000784	0	0,000189	0,000675	5,35E-05	0,000376	0,002306	0	0,001272	9,7E-05	5,02E-06	0,001445	0	3,34E-05	1,67E-06	6,69E-06	8,36E-06	0,000154	0,003259	0,000212	0	2,01E-05	0,000385		
в	0,007481	1,34E-05	0,000132	2,34E-05	0,000321	0,000658	0	0,000492	0,00187	0	0,000276	0,000675	0,000139	0,001047	0,007492	0,000257	0,000727	0,002502	0,000343	0,000744	0	4,68E-05	1,67E-06	0,000129	0,001134	2,01E-05	0,003102	0,000575	0	0	0,000304		
г	0,002541	0	3,34E-06	0	0,001072	0,00266	0	0,00167	0,000784	0	8,86E-05	0,001809	1,34E-05	0,000363	0,000947	1,67E-06	0,001522	1,1E-05	2,51E-05	0,000574	0	0	0	5,93E-05	1,67E-06	0	0	0	0	0	0		
д	0,000881	8,36E-06	0,001088	1,34E-05	1,34E-05	1,34E-05	6,19E-06	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134	0,000134		
е	0,001155	0,001025	0,001481	0,00349	0,003342	0,001535	0,000343	0,000165	0,000237	0,002205	0,001707	0,000674	0,004337	0,000444	0,000231	0,00141	0,007705	0,004834	0,000617	0,000119	0,00014	0,000472	0,000393	0,001262	0,000774	0,000689	0	0	0,000212	0,00020	0,00020		
ж	0,001567	2,84E-05	0	0,167E-05	0,000853	0,00352	1,17E-05	0	0,001344	0	0,000167	4,85E-05	1E-05	0,000749	6,35E-05	0	5,02E-06	2,51E-05	3,34E-06	0,000221	0	0	0	1,81E-05	0	0	0	0,17E-05	0	0	0		
з	0,00649	0,002022	0,001296	0,00324	0,001077	0,000747	0,001004	0,001004	0,000584	0	0,001344	0,00041	0,000279	0,001642	0,000771	4,68E-05	0,000398	1,5E-05	5,02E-06	0,000505	0	0	0	1,67E-05	3,34E-06	0	0,000455	0,000181	1,167E-06	0,000237	0,000237		
и	0,000273	0,000669	0,003282	0,000585	0,001766	0,001966	0,000294	0,002607	0,000574	0,001926	0,00318	0,007295	0,003013	0,003851	0,00048	0,000353	0,001316	0,000443	0,000461	0,000144	9,2E-05	0,001712	0,001155	0,002035	0,00059	0,000234	0	0	3,34E-06	0,000291	0,001179		
й	0	1,67E-05	0	0	1,34E-05	0,00012	1,67E-06	0	0	0	0,000155	1,84E-05	4,68E-05	0,000548	1,34E-05	1,67E-06	0	0,000451	0,00046	0	3,34E-06	0	9,86E-05	0,000197	0,000204	2,17E-05	0	0	0	0	1,67E-05		
к	0,000841	1,67E-06	0,000462	3,34E-06	0	0,00103	5,68E-05	1,84E-05	0,003431	0	2,68E-05	0,00096	3,34E-06	0,000083	0,010004	0	0,002503	0,000169	0,000833	0,002595	0	3,34E-06	3,01E-05	3,34E-06	1,34E-05	0	0	0	0	0	0		
л	0,00895	9,36E-05	2,84E-05	0,000135	7,86E-05	0,00529	0,000252	3,51E-05	0,007414	0	0,000472	0,00058	6,35E-05	0,000017	0,007384	8,66E-05	0	0,002336	0,000232	0,001547	1,34E-05	6,69E-06	0	0,000157	1,67E-05	1,67E-06	0,00102	0,00163	1,67E-06	0,000804	0,001568		
м	0,004254	1,34E-05	3,51E-05	9,03E-05	8,36E-06	0,003697	0	0	0,003332	0	0,000129	0,000219	0,00001	0,001518	0,003453	0,000206	0,000112	0,000321	9,03E-05	0,00205	5,02E-06	0	8,36E-06	3,18E-05	1,34E-05	3,34E-06	0,000829	5,35E-05	3,34E-06	0	0,000614		
н	0,010878	6,69E-06	0	0,000112	0,001015	0,001045	1,34E-05	6,69E-05	0,000853	0	0,000308	0	0	0,000387	0,011559	0	4,68E-05	0,00059	0,00097	0,000439	8,19E-05	0	0,000361	1,19E-05	5,02E-06	0,000206	0,003797	0,001094	0	0,000187	0,01662		
о	1,34E-05	0,003545	0,002655	0,005163	0,005349	0,001817	0,001859	0,001761	0,000788	0,004533	0,003035	0,0010452	0,000737	0,00638	0,000231	0,001612	0,007333	0,000623	0,007705	6,69E-05	0,000365	0,000575	0,00244	0,002032	0,000918	0,000263	0	0	0,000167	0,000475	0,000591		
п	0,001981	0	0	0	0	0,0002525	0	0	0,001491	0	0,000107	0,000165	0	0,000303	0,019962	3,34E-06	0,000721	0,0014E-05	0,000185	0,00095	0	0	5,02E-06	2,51E-05	2,01E-05	0	0	0,000393	0,001004	0	0,51E-05	0,000451	
р	0,00896	8,03E-05	0,000579	0,001204	0,003036	0,006666	0,000371	5,52E-05	0,006618	0	0,000472	0,000304	0,000234	0,001137	0,009424	7,51E-05	7,02E-05	0,000197	0,001099	0,002963	5,02E-06	0,000251	7,86E-05	0,000219	0,000532	4,85E-05	0,001587	0,000741	0	0,000298	0,012028		
с	0,001981	8,53E-05	0,001757	1,38E-05	0,000298	0,003965	4,51E-05	2,17E-05	0,00214	0	0,004626	0,002934	0,000706	0,001025	0,003095	0,001916	0,002007	0,001092	0,01089	0,000891	2,51E-05	0,000157	8,36E-05	0,000378	0,000159	0,000534	3,34E-06	0,000381	0,000415	1,67E-06	0,000112	0,000457	
т	0,006628	2,84E-05	0,000232	3,34E-06	0,000181	0,000785	0	5,02E-06	0,000408	0	0,000858	0,000252	3,34E-05	0,001483	0,012744	9,7E-05	0,003391	0,000395	6,35E-05	0,002339	6,69E-05	1,5E-05	2,01E-05	0,000249	6,69E-06	0,001674	0,000536	0	8,36E-05	0,00005			
у	0,000012	0,000756	0,000853	0,000105	0,000214	0,000679	0,00142	0,000331	2,51E-05	0,000174	0,000117	0,002558	0,001105	0,000498	5,02E-06	0,000734	0,00115	0,001431	0,001936	5,02E-06	0,000232	1,00E-05	1E-05	0,000893	0,000823	0,000298	0	0	1,84E-05	0,001182	0,00105		
ф	0,000012	1,67E-06	0	0	0	0,000105	0	0	0,000105	0	0	0,000105	0	0,000105	0	1,67E-06	0	0,000105	2,51E-05	6,19E-05	1,34E-05	0	0	0	0	0	0	0	0	0	0		
х	0,000674	1,67E-06	0,000197	6,69E-06	0	0	7,51E-05	0	0,000333	0	8,36E-06	0,000199	8,66E-05	0,000395	0,000215	0	0,000234	5,52E-05	2,51E-05	0,000219	0	3,34E-06	0	2,84E-05	0	0	0	8,36E-06	3,34E-06	0	0		
ц	0,000751	1,67E-06	7,51E-05	0	0	0,000988	0	0	0,000398	0	1,67E-05	6,69E-06	0	0	0,000306	0	4,51E-05	0	0	3,34E-06	0,000184	0	0	0	0	0	0	0,000185	0	0	0		
ч	0,002667	0	1,67E-06	0	0	0,000504	0	0	0,001843	0	0,000584	3,01E-05	5,02E-06	0,00025	7,52E-05	0	7,86E-05	0	0,003622	0,000568	0	0	0	0	0	0	0	6,19E-05	0	0,000303	0		
ш	0,001212	0	7,69E-05	0	0	0,002446	0	0	0,002145	0	0,000465	0,00051	8,36E-06	0,000453	0,000288	5,68E-05	0,000114	0,000461	0	0	0	0	0	0	0	0	0	0	0	0,0004	0	0	
щ	0,000525	0	0	0	0	0,001481	0	0	0,001222	0	0	0	0	0	6,52E-05	5,02E-06	0	3,34E-06	0	0	0	0	0	0	0	0	0	0	0	2,51E-05	0	0	0
ы	0	0,000304	0,00099	0,00012	0,000117	0,001063	8,69E-05	8,86E-05	0,000466	0,002118	0,000239	0,002274	0,001473	0,000181	0	0,000162	0,000299	0,000789	0,00065	0	0	0,000995	1,84E-05	0,000181	0,000594	8,36E-06	0	0	0	0	1,84E-05	0	
ь	0	4,68E-05	4,68E-05	0,000129	0,001052	1,67E-06	0,000144	7,02E-05	0	0,001075	0	0,000268	0,001481	8,36E-06	1,34E-05	1,67E-06	0,000836	0,000291	0	1,5E-05	0	0,00014	7,02E-05	0,000383	1,17E-05	0	0	5,02E-06	0,000391	0,000493	0	0	
э	0	0	0	5,02E-06	0	0	0	0	0	1,17E-05	3,68E-05	2,51E-05	1,17E-05	2,01E-05	0	5,02E-06	1,34E-05	1,67E-05	0,000268	0	0	8,36E-06	2,84E-05	0	0	0	0	0	0	0	0	0	
ю	0	0,000199	2,01E-05	1E-05	0,000278	0	6,69E-06	1,67E-06	0	0	4,35E-05	3,18E-05	7,19E-05	4,68E-05	0	1,67E-06	5,52E-05	0,000202	0,000343	0	1,67E-06	7,86E-05	2,01E-05	0,002135	3,85E-05	0,000558	0	0	6,66E-05	1,67E-06	0	0	
я	0	2,68E-05	0,000033	0,000115	0,000655	0,00011	0,000237	0,000308	3,18E-05	5,68E-05	0,00013	0,00065	0,000366	0,00065	0	6,86E-05	0,00011	0,00001	0,000473	1,34E-05	0,000365	0,000176	5,02E-05	0,000186	2,84E-05	0,00044	0	0	0,000169	0,00015	0	0	

×

Произвольная часть текста:
м_я_бы_по

Использованные буквы:

Порядок n-граммы:
5 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ:

Символ по счету:

Номер эксперимента: 51

Поле ввода символов:

Продолжить

Другой

Неравенство для энтропии:
 $2,76330013614539 < H < 3,46965350016743$

Двоичная таблица угаданных символов:
0000000000000000000100000000000 ^
1000000000000000000000000000000
0000000000010000000000000000000
0000000000000000000000000000001
0000100000000000000000000000000 v

Вероятности:
q[1] = 0,36
q[2] = 0,14
q[3] = 0,04
q[4] = 0,04
q[5] = 0,02
q[6] = 0,02
q[7] = 0,02
q[8] = 0,02
q[9] = 0
q[10] = 0,02
q[11] = 0,02
q[12] = 0,04
q[13] = 0
q[14] = 0
q[15] = 0
q[16] = 0
q[17] = 0,04
q[18] = 0,02
q[19] = 0
q[20] = 0,04
q[21] = 0,02
q[22] = 0,04
q[23] = 0
q[24] = 0,02
q[25] = 0,02
q[26] = 0
q[27] = 0
q[28] = 0
q[29] = 0,02
q[30] = 0
q[31] = 0
q[32] = 0,04

Строка состояния:

✕

Произвольная часть текста:
какой_смысл_заявят

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ:

Символ по счету:

Номер эксперимента: 54

Поле ввода символов:
Продолжить Другой

Неравенство для энтропии:
2,18060339754625 < H < 2,83729433290204

Двоичная таблица угаданных символов:
10000000000000000000000000000000 ^
10000000000000000000000000000000
00000000000001000000000000000000
00000000000000000100000000000000
00000000000000000000000000010000 v

Вероятности:
q[1] = 0.4339622
q[2] = 0.1132075
q[3] = 0.1320754
q[4] = 0.0377358
q[5] = 0.0377358
q[6] = 0.0188679
q[7] = 0
q[8] = 0
q[9] = 0
q[10] = 0
q[11] = 0.037735
q[12] = 0
q[13] = 0.037735
q[14] = 0
q[15] = 0
q[16] = 0
q[17] = 0
q[18] = 0.018867
q[19] = 0.037735
q[20] = 0.037735
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0
q[25] = 0.018867
q[26] = 0
q[27] = 0
q[28] = 0.037735
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0

Строка состояния:

Произвольная часть текста:
и_по_тому_вопросу_сколько_же

Использованные буквы:

Порядок n-граммы:

- ☐ 5 символов
- ☐ 10 символов
- ☐ 15 символов
- ☐ 20 символов
- ☒ 25 символов
- ☐ 35 символов
- ☐ 40 символов
- ☐ 45 символов
- ☐ 50 символов

Введенный символ:

Символ по счету:

Номер эксперимента: 52

Поле ввода символов:

Продолжить Другой

Неравенство для энтропии:
 $H < 3,1971573835724$

Двоичная таблица угаданных символов:

00000010000000000000000000000000	▲
000000000000000000000000100000000000	
0100000000000000000000000000000000	
1000000000000000000000000000000000	
1000000000000000000000000000000000	▼

Вероятности:

q[1] = 0,4117647
q[2] = 0,1568627
q[3] = 0,0392156
q[4] = 0,0392156
q[5] = 0
q[6] = 0,0196078
q[7] = 0,0196078
q[8] = 0
q[9] = 0
q[10] = 0,019607
q[11] = 0
q[12] = 0
q[13] = 0,039215
q[14] = 0,019607
q[15] = 0
q[16] = 0
q[17] = 0,039215
q[18] = 0,019607
q[19] = 0
q[20] = 0
q[21] = 0,019607
q[22] = 0,019607
q[23] = 0
q[24] = 0,019607
q[25] = 0
q[26] = 0,039215
q[27] = 0,019607
q[28] = 0,019607
q[29] = 0,019607
q[30] = 0
q[31] = 0,019607
q[32] = 0

Строка состояния:

Оцінка надлишковості мови:

H	R
H1 with whitespaces	0,1258
H1 without whitespaces	0,101628
H2 with whitespaces	0,41922
H2 without whitespaces	0,323822
H10	0,448>R>0,308
H20	0,564>R>0,434
H30	0,499>R>0,361

Висновки:

Під час виконання даної лабораторної роботи я навчився вимірювати частоти символів та біграм у тексті, визначати ентропію. Також, я навчився визначати надлишковість мови.